日 本 国 庁 JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日 Date of Application:

2003年 1月14日

出願番 Application Number:

特願2003-005237

[ST.10/C]:

[JP2003-0052371

Ш 人 Applicant(s):

株式会社日立製作所

U.S. Appln Filed 7-14-03
Inventor: K. Wakayama etas
mattingly Stanger & Malur

Oocket H-1100

2003年 4月25日

特許庁長官



【書類名】 特許願

【整理番号】 H02016091A

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/56

【発明者】

【住所又は居所】 東京都国分寺市東恋ケ窪一丁目280番地 株式会社日

立製作所中央研究所内

【氏名】 若山 浩二

【発明者】

【住所又は居所】 東京都国分寺市東恋ケ窪一丁目280番地 株式会社日

立製作所中央研究所内

【氏名】 森脇 紀彦

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【電話番号】 03-3212-1111

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】

明細書

【発明の名称】 パケット通信方法およびパケット通信装置

【特許請求の範囲】

【請求項1】

パケットを送受信する複数のインタフェースと、該複数のインタフェースに接続されたスイッチと、該スイッチに接続された統計情報収集処理部と、前記パケットに付与されたヘッダ情報を解析する手段と、前記インタフェースを介して送信または受信されるパケットの量を計数する手段とを備え、

前記統計情報収集処理部は、前記解析されたヘッダ情報とパケット量とから前 記複数のインタフェースで受信されるパケット量を予測し、

該予測されたパケット量に基づき、パケットを送信するインタフェースを選択 することを特徴とするパケット通信装置。

【請求項2】

請求項1に記載のパケット通信装置において、前記ヘッダ情報を解析する手段 と前記パケットの量を計数する手段とは、前記統計情報収集処理部に含まれるこ とを特徴とするパケット通信装置。

【請求項3】

請求項1に記載のパケット通信装置において、前記複数のインタフェースに接続され、かつ前記ヘッダ情報を解析する手段と前記パケットの量を計数する手段とを備えた回線対応部を有することを特徴とするパケット通信装置。

【請求項4】

請求項1に記載のパケット通信装置において、前記インタフェースと前記統計 情報収集処理部とを直接接続するバスを備えたことを特徴とするパケット通信装 置。

【請求項5】

請求項1に記載のパケット通信装置において、前記パケットを送受信するインタフェースは、該インタフェースが送受信する少なくとも一つ以上のパケットに付与されたヘッダ部の少なくとも一部をフレームに格納する手段を有することを特徴とするパケット通信装置。

【請求項6】

請求項5に記載のパケット通信装置において、前記フレームに多重化される複数のパケットのヘッダ部の大きさは全て等しいことを特徴とするパケット通信装置。

【請求項7】

請求項5に記載のパケット通信装置において、前記多重化する手段は、前記へ ッダ情報を複数のパケットから、それぞれのパケットに付与されるヘッダに設定 された該パケットの種別を表す情報に応じて、切り出すヘッダ部分の長さを決定 し、一つのフレームに多重化することを特徴とするパケット通信装置。

【請求項8】

パケット通信装置。

請求項1に記載のパケット通信装置において、前記統計情報収集処理部を複数 備えることを特徴とするパケット通信装置。

【請求項9】

請求項1に記載のパケット通信装置において、前記負荷分散装置に接続された 拡張処理部を有し、当該拡張処理部は、受信パケットの転送が実行されるレイヤ よりも上位のレイヤで実行される処理を行なうことを特徴とするパケット通信装 置。

【請求項10】

請求項3に記載のパケット通信装置において、前記回線対応部に設けられ、かつ受信パケットのヘッダ情報とパケットの転送先との対応関係を記述するテーブルを備えることを特徴とするパケット通信装置。

【請求項11】

請求項10に記載のパケット通信装置において、前記予測されたパケット量に基づき、前記テーブルを更新する手段を備えることを特徴とするパケット通信装置。

【請求項12】

パケットを送受信する複数のインタフェースと、パケットの処理手段とを備え たパケット通信装置に用いられるパケット通信方法において、 前記インタフェースを介してパケットを受信するステップと、

該受信したパケットの数を計数するステップと、

前記計数したパケットの数に基づき前記複数のインタフェースの1に将来到達 するパケットの数を予測するステップと、

該予測されたパケットの数に基づき送信パケットを伝送するインタフェースを 選択するステップとを有するパケット通信方法。

【請求項13】

請求項12に記載のパケット通信方法において、複数の前記受信パケットのヘッダ情報を一つに多重化するステップを有することを特徴とするパケット通信方法。

【請求項14】

請求項13に記載のパケット通信方法において、受信パケットから前記ヘッダ 部を固定長分だけ切り出すステップとを有することを特徴とするパケット通信方 法。

【請求項15】

請求項14に記載のパケット通信方法において、受信パケットのヘッダ部をそれぞれのパケットに付与されるヘッダに設定された該パケットの種別を表す情報に応じた大きさだけ切り出すステップを有することを特徴とするパケット通信方法。

【請求項16】

パケットを送受信する複数のインタフェースと、該複数のインタフェースに接続されたスイッチと、該スイッチに接続された統計情報収集処理部と、前記パケットに付与されたヘッダ情報を解析する手段と、前記インタフェースを介して送信または受信されるパケットの量を計数する手段とを備え、

前記統計情報収集処理部は、前記計数されたパケット量に基づき、パケットを 送信するインタフェースを選択することを特徴とするパケット通信装置。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】

本発明は、パケット通信装置に関するものであり、特に統計情報収集方式、ならびに負荷分散方式に関する。

[0002]

【従来の技術】

電子商取引や映像配信等、インターネットを用いた通信サービスの高度化により、トラヒック量の増加と通信速度の高速化は急激に進行している。これに伴って、ネットワーク内においてルータやサーバ等の通信装置の処理負荷が一層大きくなってきている。このため、通信装置の処理能力の向上が急務となっている。たとえば、WWW (World Wide Web) サービスにおいては、特定サーバへのアクセスが増加し、単一のサーバで処理をまかなうには、能力が不十分な場合が生じている。

[0003]

また、インターネットを用いた仮想プライベート網(Virtual Private Network、VPN)サービスを実現するためには、送受信するパケット単位に行う高位レイヤ処理を通信速度に即して行うことが必要である。高位レイヤ処理としては、暗号化、ファイアウォール処理等があげられる。

[0004]

パケットの転送方式に関しても、種々の方式が提案されている。例えば、IEEE 802.3で規定されるEthernet (登録商標) フレームによる転送、IEEE802.1Qで規定されるタグVLANによる転送、IETF RFC3032で規定されるMPLS (Multi-Protocol Label Switching) による転送、IEEE802.17委員会で標準化作業中のRPR (Resilient Packet Ring) による転送がある。これらは、いずれもEthernetを用いたパケット転送の方式であるが、他にも通信方式により種々のパケット転送方式がある。

[0005]

通信装置の処理能力を向上する手法の一つとして、負荷分散技術がある。負荷分散技術は、同じ機能を持つ装置を複数台設け、これらの装置間で処理を分担することにより、各単一装置における処理負荷を低減するとともに、全体として処理能力を向上を図る方法である。

[0006]

負荷分散技術を用いることによって、処理能力の向上の他、耐障害性の向上を も図ることが可能になる。すなわち、一台の装置に故障が発生した際においても 、残りの装置が処理を引き継ぐことにより、運用を継続することが可能になる。

[0007]

負荷分散には、装置内における処理負荷を分散することを目的とした負荷分散 と、該装置に接続された他の装置(たとえば、サーバ)の処理負荷を分散するこ とを目的とした負荷分散がある。

[0008]

負荷分散を実現するためには、負荷分散の対象となる複数の装置に対して、適切にパケットを振り分けて転送することが必要である。

[0009]

パケットの振り分けを実現する一つめの方式として、ラウンドロビンによる方法がある。ラウンドロビンによる方法とは、パケットを、受信した順序にしたがって、互いに異なる装置に対して振り分けていく方法である。

[0010]

パケットの振り分けを実現する二つめの方式として、ヘッダ情報に対するハッシュ値にしたがって、パケットを振り分ける方法がある。ハッシュ値による方法とは、フローを表すヘッダ情報、たとえば、パケットの送信元 I Pアドレスと宛先 I Pアドレスの組を引数として、ハッシュ関数を計算し、得られた値 (ハッシュ値) にしたがって、パケットの振り分け先を決定する方法である。

[0011]

パケットの振り分けを実現する三つめの方式として、フィルタリングテーブルを用いる方法がある。フィルタリングテーブルを用いる方式とは、あらかじめ、フローと、該フローのパケットに対する振り分け先を対応付けるテーブルを用意し、パケット受信時には、パケットのヘッダに示されているフローを表す情報を検索キーとして前記テーブルを検索することによって、パケットの振り分け先を決定する方法である。

[0012]

【発明が解決しようとする課題】

ラウンドロビンによるパケット振り分けでは、パケットの到着順にしたがって振り分けを行うため、振り分け先に対して比較的均等に分散させることが可能である。しかしながら、パケット転送による遅延時間等、処理に要する時間の差が生じうるため、同じフロー(たとえば、送信元アドレスと宛先アドレスの組)において、負荷分散対象となる装置が受信したパケットの順序と、前記の負荷分散対象が送信するパケットの順序が異なる可能性があることが問題である。前記問題を解決する方法としては、パケット受信時に、パケットに付与する内部ヘッダにおいて、シーケンス番号を付与し、パケット送信時には、シーケンス番号の順序にしたがってパケットを送信する方法がある。しかしながら、フロー数が増加すると、シーケンス番号管理の処理が複雑化することが問題となる。

[0013]

ハッシュ値による方法では、同一フローに対しては、ハッシュ値の値が同一になるため、同じフローに対して、受信したパケットの順番にしたがって、パケットが送信されることになる。しかしながら、ハッシュ値による方法では、特定のフローが多い場合には、必ずしも均等にパケットを振り分けられない可能性のあることが問題である。

[0014]

フィルタリングテーブルによる方法では、あらかじめフローを登録しておくことが必要である。また、フローごとのトラヒックパターンの変動に伴って、パケット振り分け先が偏る可能性があることが問題である。

[0015]

そこで、本発明の目的は、実際のトラヒックの状況に応じて、均等な振り分け を実現することにある。

[0016]

【課題を解決するための手段】

本発明においては、実際のトラヒックの状況に応じた均等な振り分けを実現するため、パケット量の統計情報を利用する。装置のトラヒック状況は、インタフェースを通過するパケットのパケット量を計数することにより把握できる。また

、統計情報の収集を、装置内でパケットの転送処理を行う際、パケット数、バイト数の積算、およびパケットに付与されたヘッダ情報を記録することにより収集できる。

[0017]

また、パケットに付与されたヘッダ情報のフォーマットは、パケット転送方式により、異なるため、ルータにおいては、パケット転送方式に応じた統計情報収集処理が必要となる。更に、収集すべき統計情報の項目は、装置の適用箇所等に応じて異なる。したがって、多様なパケット転送方式に対応した統計情報を収集するためには、ルータにおける処理が複雑化することが問題になる。

[0018]

そこで、本発明においては、統計情報を収集することを目的とした専用の機能 ブロックを設ける。パケットの送受信を行う回線インタフェースは、パケットに 付与されたヘッダ情報を、前記の統計情報収集処理を行う機能ブロックへ転送す る。これによって、パケット転送処理と統計情報収集処理を独立した機能ブロッ クにおいて行う。

[0019]

パケット転送をこのように行なうことで、多様なパケット転送方式に対応した 統計情報収集処理を実現可能となり、高速、高機能なパケット通信装置が可能と なる。

[0020]

具体的には、パケット通信装置内に、パケットの負荷分散を行なう手段と、インタフェースを通過するパケットの計数手段と、計数されたパケットの統計処理を行なう統計情報収集処理部を設けて、インタフェースで受信されるパケット量を予測する。負荷分散手段は、予測されたインタフェースのトラヒックを元に負荷分散制御を行なう。

[0021]

【発明の実施の形態】

以下、図面を用いて発明の実施の形態を説明する。

[0022]

図6は、本発明によるパケット通信装置のネットワーク構成の一例を示すものである。

[0023]

図6の例においては、ネットワーク100-1から100-11は、パケット通信装置1-1から1-16、サーバ2-1から2-4によって構成されている。ネットワーク100-1、100-2においては、パケットはRPRにより転送される。ネットワーク100-3におけるパケット転送方式は、Ethernetにより転送される。ネットワーク100-4、100-5、100-8においては、パケットはVLANにより転送される。ネットワーク100-6、100-9においては、パケットはATM(Asynchronous Transfer Mode)により転送される。ネットワーク100-7、100-10、100-11においては、パケットはMPLSにより転送される。

[0024]

図1は、本発明によるパケット通信装置の構成の一例を示すものである。 パケット通信装置1は、物理回線の終端、パケット転送処理を行う回線対応部1 1、パケットをスイッチングするスイッチ12、装置制御を行う制御部13、暗 号化処理等、高位レイヤ処理を行う拡張処理部14、統計情報の収集処理を行う 統計情報収集処理部15、ならびに、制御部13と、前記回線対応部11、スイッチ12、拡張処理部14とを接続し、装置内部の制御用信号の転送を行う装置 内バス16から構成される。

[0025]

回線対応部11と統計情報収集処理部15の間で転送する信号の転送方法として,スイッチ12を経由する方法,または,バス16を経由する方法が挙げられる。どちらの方法を用いるのが適切かは,スイッチ12のスイッチング容量や,バス16の転送能力,または,装置内の制御信号の情報量等により異なるが、バスを経由して信号を転送することにより、制御信号がスイッチを経由せずに流れるため、パケット通信装置1が送受信するパケットのトラヒックが過負荷な状態においても、制御信号転送によるスイッチング能力の低下を防ぐことができる。

[0026]

図1を用いて、パケット通信装置1におけるパケット処理の流れについて説明する。回線対応部11は、受信したパケットのヘッダ情報を検索することによって、該パケットを通信装置内におけるどの回線対応部11、または拡張処理部14へ転送すべきかを決定し、該パケットに、該パケットの転送先を示す装置内ヘッダを付与してスイッチ12へ転送する。スイッチ12では、前記パケットに付与された装置内ヘッダの情報を参照することによって、適切な回線対応部11、または拡張処理部14に対して該パケットをスイッチングする。回線対応部11においては、スイッチ12からスイッチングされたパケットに対して、物理回線へ送出するためのカプセル化処理を行い、物理回線へ該パケットを送出する。また、拡張処理部14へパケットがスイッチングされた場合においては、拡張処理部14は該パケットに対して、暗号化処理等の高位レイヤ処理を行う。

[0027]

本発明によるパケット通信装置が行う負荷分散処理の一つめの例は、図21で示されるような、回線対応部11-1が受信したパケットに対して高位レイヤ処理を行う際、処理負荷を拡張処理部14-1から拡張処理部14-mが分散するために、回線対応部11-1から、パケット30-1から30-mを、それぞれ拡張処理部14-1から拡張処理部14-mに対して振り分ける処理である。ここで、本実施例では、パケットを受信し、受信パケットに付与された宛先アドレスを元に、パケットの転送先を決定し、パケットを送信する処理を基本パケット転送処理と定義する。基本パケット転送処理が扱うレイヤよりも上位のレイヤが扱う処理を高位レイヤ処理と定義する。

[0028]

本発明によるパケット通信装置が行う負荷分散処理の二つめの例は、図6のパケット通信装置1-7によるサーバ2-1またはサーバ2-2に対するパケット振分処理、または、パケット通信装置1-8によるサーバ2-3またはサーバ2-4に対するパケット振分処理である。負荷分散処理の二つめの例においては、該サーバに接続する回線対応部に対してパケットを振り分ける処理と同義である

[0029]

図2は、図1で示したパケット通信装置の回線対応部11の構成の一例を示す ものである。

[0030]

回線対応部11は、受信パケットの物理レイヤ終端を行うレイヤ1受信処理部112、送信パケットの物理レイヤ終端処理を行うレイヤ1送信処理部113、受信パケットをバッファリングする受信バッファ114、送信パケットをバッファリングする送信バッファ115、パケットの転送先の決定、パケットへのカプセル化処理、すなわち、転送時にパケットにヘッダを付与するパケット処理エンジン116、パケット処理エンジン116が、パケットに付与されたヘッダ情報と、該パケットに対する処理の対応関係に関する情報を格納する検索テーブル117、スイッチとのインタフェースであるスイッチインタフェース118、制御部との通信インタフェースである制御部インタフェース119、送受信パケットに付与されているヘッダ情報を格納するヘッダ格納バッファ120、プロセッサ121、メモリ122から構成される。検索テーブル117はメモリ122に格納される場合もある。

[0031]

図2を用いて回線対応部11における受信パケットに対する処理方法について説明する。レイヤ1受信処理部112において、光一電気変換処理等の物理レイヤ処理を行い、受信パケットバッファ114でパケットをバッファリングする。パケット処理エンジン116では、受信パケットのヘッダ情報をキーとして、検索テーブル117を検索し、パケットの転送先決定、ヘッダカプセル化の処理を行う。該パケットを、別の回線対応部、または拡張処理部へ転送する際には、スイッチインタフェース118-1からスイッチ12へ送出する。また、スイッチインタフェース118-1からスイッチ12よりパケットを受信した際には、最初に、パケットを送信パケットバッファ115にバッファリングする。次に、パケット処理エンジン116は、該パケットの先頭に付与された内部ヘッダの情報をキーに検索テーブル117を検索し、物理回線からパケットを送出するためのカプセル化処理を行い。レイヤ1送信処理部113において、電気一光変換処理等の物理レイヤ処理を行い、物理回線に対してパケットを送出する。

[0032]

また、パケット処理エンジン116は、パケットに付与されたヘッダ情報を切り出してヘッダ格納バッファ120にバッファリングし、同時に、該パケットの長さを計数する。

[0033]

図7は、パケットを、受信側回線対応部11からスイッチ12を経由して送信側回線対応部11へ転送する際のフォーマットの一例を示すものである。回線対応部11は、受信側回線対応部11から送信側回線対応部11へ転送するパケット301の先頭に、装置内でパケットを転送するために必要な情報を設定するための装置内へッダ310を付与する。装置内へッダ310には、パケットを、パケット通信装置1から送信するために必要となる情報、すなわち、回線対応部番号、送信側の物理ポート番号、送信側物理ポートの回線種別、パケット通信装置1からの転送先装置のIP(Internet Protocol)アドレスを設定するパケット出力情報311、受信側回線対応部11に関する情報、すなわち、受信側回線対応部の番号、パケットを受信した物理ポートの番号、パケットを受信した物理ポートの回線種別を設定するパケット入力情報312、該パケットが、ユーザデータパケットであるか、装置内制御信号のパケットであるか等、パケットの種別を表すパケット識別子313、装置内におけるパケットの転送優先度314が設定される。パケット出力情報311、優先度314への設定値は、回線対応部11における検索テーブル117の検索結果により決定する。

[0034]

図3は、検索テーブル117の構成の一例を示す図である。検索テーブル117は、検索キー1171と検索結果1172の対応を関係を表すエントリ1173を設ける。図3の例においては、検索キーとして、パケットの送信元IPアドレスとパケットの宛先IPアドレスの組を用いている。

[0035]

図4は、本発明によるパケット通信装置における統計情報収集処理部の構成の一例を示す図である。統計情報収集処理部15は、回線対応部から転送されたヘッダ転送フレームを格納するデータバッファ151、統計情報を収集するために

ヘッダ情報読み取り処理を行うヘッダ情報読み取り部152(ヘッダ情報の解析手段)、パケット通信装置1において受信したパケットの、フローごとのパケットの量、例えばパケット数、バイト数、ビット数を計算するデータ加算部153(パケット量の計数手段)、データ加算部で計算した統計情報を格納する統計テーブル154、統計情報収集処理部15の制御、テーブル更新処理を行うためのプロセッサ155、プロセッサのプログラム領域であるメモリ156、スイッチとのインタフェースであるスイッチインタフェース157、制御部との通信インタフェースである制御部インタフェース158等により構成される。

[0036]

図8は、回線対応部からヘッダ情報を統計情報収集処理部に対して転送する方法の一例として、複数パケットのヘッダ情報を多重化したフレームを回線対応部において生成し、統計情報収集処理部へ転送する際のヘッダ転送フレームのフォーマットの例を示すものである。ヘッダ転送フレーム35は、図7で示した装置内ヘッダ310とヘッダ領域37から構成される。

[0037]

回線対応部11は、装置内ヘッダ310のパケット出力情報を用いて、フレーム35の転送先として、統計情報収集処理部15を指定する。

[0038]

ヘッダ領域37は、該ヘッダが付与されていたパケットの入力ポート371、 該ヘッダの長さ372、該ヘッダが付与されていたパケットの長さ373を設定 するパケット情報部370、およびヘッダ情報374から構成される。

[0039]

ヘッダ情報374の長さは、物理回線におけるパケット転送方式に関わらず、 固定長にしてもよい。また、ヘッダ情報374の長さは、物理回線におけるパケット転送方式に応じて可変長としてもよい。 一つのヘッダ転送フレーム35には、N個のパケットのヘッダ領域37-1から37-nが格納される。

[0040]

図8の例のように,一つのヘッダ転送フレームに複数個のパケットのヘッダ領域を格納して転送、つまりヘッダ情報を多重化して転送することにより、回線対

応部 1 1 から統計情報収集処理部 1 5 に対してヘッダ情報を転送することに伴う オーバヘッドを削減することが可能である。

[0041]

図4を用いて統計情報収集処理部における統計情報収集処理を説明する。

[0042]

回線対応部11から転送されたヘッダ転送フレーム35は、スイッチインタフェース157-1を経由してデータバッファ151に格納される。ヘッダ情報読み取り部152は、データバッファ157-1から、ヘッダ転送フレーム35内のヘッダ領域37-1から37-Nの、個々のヘッダ領域を切り出す。さらに、ヘッダ情報読み取り部152は、それぞれのヘッダ領域37-1から37-Nについて、それぞれパケット情報部370-1から370-Nに設定されているパケット長、および、それぞれヘッダ情報374-1から374-Nに設定されている送信元アドレス、宛先アドレス等、統計情報として収集すべきフロー情報を読み出す。データ加算部は153、ヘッダ情報読み取り部で取り出したフロー毎のパケット数、バイト数、ビット数を加算するためのレジスタを保持する。データ加算部153のレジスタに書き込まれたパケット数、バイト数、ビット数は、一定時間ごとに統計テーブル154に書き込む。

[0043]

図5は、統計情報収集処理部15に設ける統計テーブル154の構成の一例を示すものである。統計テーブル154には、フローを設定するフィールド1541と、前記フローにおける統計情報を記録するフィールド1542から構成されるエントリ1543を設ける。図5の統計テーブル154の例では、フロー1541として、パケットの送信元IPアドレスと宛先IPアドレスの組としている。また、収集する統計情報の項目として、該フローの入力回線対応部の識別情報15421、入力物理ポートの識別情報15422、および、パケット数の総計15423、パケットのバイト数の総計15424としている。

[0044]

なお、ここまでの実施の形態では、ヘッダ情報は、図22におけるヘッダ転送 フレーム35-1、のように、回線対応部11からスイッチ12を経由して、統 計情報収集処理部15へ転送する場合の統計情報収集処理部の構成の一例について説明した。ヘッダ情報は、図22におけるフレーム35-2のように、回線対応部11から内部バス16を経由し、統計情報収集処理部へ15転送することも可能である。ヘッダ情報を、回線対応部11から装置内バス16を経由し、統計情報収集処理部へ15転送する際には、統計情報収集処理部は制御部インタフェース158を経由して、装置内バス16との間でヘッダ情報を送受信する。

[0045]

図10は、図8で示したヘッダ転送フレーム35を用いて統計情報収集処理部 15ヘヘッダ情報を転送する際の、回線対応部11における処理のフローチャートを示すものである。

[0046]

回線対応部11が物理回線からパケットを受信(ステップ5010)すると、パケット処理エンジン116は、受信パケットバッファ114にバッファリングした該パケットのヘッダ部分を切り出し、ヘッダ格納バッファ120に格納する(ステップ5020)。パケット処理エンジン116は、パケット処理エンジンが処理したパケット数を加算するパケットカウンタを保持する。パケット処理エンジン116は、パケットカウンタの値Pnを1増やす(ステップ5030)。このとき、パケットカウンタの値Pnが、あらかじめ設定された値N(Nは2以上の整数)と合致するかどうかを判定する(ステップ5040)。パケットカウントの値PnがNの場合には、図4で示したヘッダ転送フレーム35を生成し、統計情報収集処理部15に対して転送する(ステップ5050)。同時にパケットカウンタの値Pnをリセットし(ステップ5060)、処理を終了する(ステップ5070)。ステップ5070)。

[0047]

なお、回線対応部 1 1 から統計情報収集処理部 1 5 に対して受信パケットのヘッダ情報を転送する際、図 9 に示すように、一つのフレームで単一のパケットのヘッダ情報を転送してもよい。一つのフレームで単一のパケットのヘッダ情報を転送する際には、図 1 0 のフローチャートにおいて、ステップ 5 0 2 0 の処理の

後、ステップ5050の処理を行うことによって、ヘッダ情報を回線対応部11 から統計情報収集処理部15へ通知する。

[0048]

ここまでの説明では、ヘッダ情報の転送処理を、パケット通信装置1が受信したパケットに対して行う場合の実施の形態を示した。ヘッダ転送処理は、パケット通信装置1が送信するパケットに対して行うことも可能である。パケット通信装置1が送信するパケットに対して、ヘッダ転送処理を行う場合には、図10のフローチャートのステップ5010において、スイッチ12から回線対応部11へ転送されたパケットを受信することになる。

[0049]

ここで、パケット通信装置 1 が処理するパケットのフォーマットの例を説明する。

[0050]

図12は、Ethernetヘッダのフォーマットを示すものである。Ethernetヘッダ 600は、宛先MACアドレス 601、送信元MACアドレス 602、カプセル化タイプを表すTypeフィールド 603 によって構成される。

[0051]

Ethernetヘッダでカプセル化されているパケットの上位プロトコルは、Ethern etヘッダ600のTypeフィールド603の値によって、識別することが可能である。

[0052]

図13はEthernetヘッダによりカプセル化されたIPパケットのフォーマットを表す。EthernetヘッダがIPパケットをカプセル化する際、Ethernetヘッダ600のTypeフィールド603には、16進数の値800が設定される。これによって、Ethernetヘッダ600の後ろには、IPヘッダ610が設定されていることを認識することを認識できる。

[0053]

図14はIEEE802.1Qで規定されるタグVLANパケットフォーマットを示す図である。Ethernetヘッダ60のTypeフィールド603に、16進数の値8100

が設定されている場合、Ethernetヘッダ600の後ろにVLANタグ620が設定されていることを認識できる。

[0054]

図15はIETF RFC3032で規定されるMPLS (Multi-Protocol Label Switch ing) パケットフォーマットを示す図である。Ethernetヘッダ60のTypeフィールド603に、16進数の値8847が設定されている場合、Ethernetヘッダ600の後ろにShimヘッダ630が設定されていることを認識できる。

[0055]

図16は、IEEE802.17委員会で標準化作業中のRPR (Resilie nt Packet Ring)のパケットフォーマットを示す図である。RPRのパケットは、RPRヘッダ630、Ethernetヘッダ600、ヘッダチェックサム631から構成される。

[0056]

次に、パケット通信装置1における統計情報収集処理の実施の形態について説明する。

[0057]

図11は、統計情報収集処理部15における統計情報収集処理の一例を示すフローチャートである。

[0058]

統計情報収集処理部15において、ヘッダ転送フレーム35を受信する(ステップ5210)ことを契機として、統計情報収集処理部15は、統計情報収集処理部15は、統計情報収集処理部15のデータ加算部153は、データ加算部153が保持するヘッダ数カウンタの値Hnを1にセットする(ステップ5220)。次に、ヘッダ転送フレーム35中の先頭から1番めのヘッダ領域37-1を切り出す(ステップ5230)。次に、ヘッダ領域37-1から、ヘッダ情報を検出することにより、該ヘッダが付与されていたパケットのフロー、たとえば、送信元アドレスと宛先アドレスの組を判定する(ステップ5240)。ここで、ヘッダ転送フレーム35の装置内ヘッダ310のパケット入力情報312、または、パケット情報部310に設定された入力ポート371の設定値によ

り、統計情報収集処理部15は、前記ヘッダ領域37-1が付与されたパケット を受信した物理回線の種類を識別することが可能である。したがって、統計情報 収集処理部15は前記ヘッダ領域37-1のヘッダフォーマットを把握すること が可能である。次に、ステップ5240において、統計テーブル154に、前記 判定したフローに関するエントリ1543が、既に存在するものであるかどうか を判定する(ステップ5250)、統計テーブル154に該フローのエントリが 存在しない場合は、エントリを新規に追加する(ステップ5260)。次に、該 ヘッダ領域37の先頭に付与されたパケット情報部370を参照する(ステップ 5270)。これによって、パケット情報部370に設定されたパケット長37 3の値から、パケットのバイト数を判定し、データ加算部153において、該フ ローのバイト数、またはビット数の値を加算する。同時に該フローのパケット数 を加算する(ステップ5280)。これらの処理を終了後、ヘッダ数カウンタの 値Hnが、フレーム35に多重されているヘッダ情報の数Nに一致するかどうか を判定する(ステップ5290)。ヘッダ数カウンタの値HnがNに一致する場 合は、処理を終了する(ステップ5300)。ヘッダ数カウンタHnの値がNに 一致しない場合には、ヘッダ数カウンタの値Hnを1増やし(ステップ5310)、ステップ5230に戻る。その後、N番めのヘッダ領域37-Nに対する処 理が終了するまで、同様の処理を繰り返す。

[0059]

次に、統計情報収集処理部15によって、回線対応部11に設ける検索テーブ ル117を更新する処理について説明する。

[0060]

図17は統計情報収集処理部における、検索テーブル更新手順の一例を示すフローチャートである。

[0061]

統計情報収集処理部15のプロセッサ155は、解析処理を起動し(ステップ5410)、テーブル更新対象となるフローを統計テーブル154から抽出する(ステップ5410)。次に、パケットの振分先の対象、すなわち、拡張処理部14-1から14-mに対して、ステップ5410で抽出したフローを割り当て

る(5420)。次にステップ5420で決定したパケットの振分先ごとのフローの割り当てにしたがって、検索テーブル117を生成する(ステップ5430)。統計情報収集処理部15は、ステップ5430で生成した検索テーブルを回線対応部11に配布する(ステップ5450)。以上により、一連の処理を終了する(ステップ5460)。

[0062]

統計情報収集処理部15による回線対応部11に対する検索テーブルの配布は、図23に示すように、図7で示した装置内のパケット転送フォーマットにより、スイッチ12、または、装置内バス16を経由して行う。

[0063]

図18は、図17のステップ5420における、それぞれの振分先に対してフローを割り当てるアルゴリズムの一例である。

[0064]

プロセッサ155は、テーブル更新対象とするフローについて、パケット数、またはバイト数の順にしたがって、フローをソーティングする(ステップ5510)。次に、振分先に対して、ステップ5510でのソーティングしたフローの昇順、または降順に1つずつフローを割り当てる(ステップ5520)。ここで、拡張処理部14-1から14-mに対して、それぞれ振分先番号1からmに対応するものとし、振分先番号1から振分先番号mの順にフローを割り当てるものとする。振分先番号mまで割り当てると、次に振分先番号mから振分先番号1に対して、ステップ5510でのソーティングしたフローの昇順、または降順に1つずつフローを割り当てる(ステップ5530)。ステップ5520、ステップ5530の割り当て処理を、全てのフローを割り当てるまで繰り返す(ステップ5540、ステップ5550)。

[0065]

なお、図17、図18のフローチャートを用いた説明では、検索テーブルの更 新処理においては、検索テーブルを更新するためのフローごとのパケット数、ま たは、バイト数として、図5で示した統計テーブルに記録された、ある一定期間 にパケット通信装置1が送信または受信したパケット数やバイト数を用いた。検 索テーブルを更新するために用いるフローごとのパケット数、または、バイト数としては、過去にパケット通信装置1が収集したフローごとのパケット数、または、バイト数を元に算出した、将来のフローごとのパケット数、またはバイト数の予測値を用いることも可能である。

[0066]

図24は、将来のフローごとのパケット数、またはバイト数の予測値を算出するために用いる統計テーブルの構成の一例を示すものである。図24に示す統計テーブルにおいては、収集した統計情報を、1542-1、1542-2、1542-30ように、それぞれ、期間 1-10、期間 1-10、用 1-10

[0067]

図25は、将来のフローごとのパケット数、またはバイト数の予測値を算出する方法の一例を説明する図である。時間 t 0 は、統計情報の収集を開始した時刻である。また、時刻 t 3 は現在時刻であり、時刻 t 4 は、次にテーブルが更新される予定の時刻である。図25の例では、時間 t 0 から t 3 の間、フロー(192.168.10.5、192.168.20.2)のパケット数は単調増加している。つまり、時間 t 0 から t 1 の期間(91-1)における、パケット通信装置 1 が受信したフロー(192.168.20.2)のパケット数が p1-p0(90-1)、時間 t 1 から t 2 の期間(91-2)における、パケット通信装置 1 が受信したフロー(192.168.10.5、192.168.20.2)のパケットがp1 の p2 の p3 の p4 の p3 における、パケット通信装置 1 が受信したフロー(192.168.10.5、192.168.20.2)のパケット数が p3 の p5 の p6 の p7 の p8 の p8 の p9 の p

 ケット数の変化を直線で近似し、その直線を時刻t4に外揮することによりパケット量の予測を行ったが、パケット量の変動がもっと複雑な場合には、種々の近似曲線、または予測手法を使用して統計予測を行っても良い。

また、過去の時点でサンプリングされたパケット量のうち、データとしていずれの時刻に収集されたパケット量を用いるかは任意に決めて良い。例えば、現在時刻 t 3 で収集されたパケット量を用いずに時刻 t 0~t 2の間で収集されたデータのみを用いて統計予測を行っても良いし、例えば、時刻 t 1 で計数されたパケット量が異常に多く、異常値と判断されるような場合には、時刻 t 0, t 2, t 3 で収集されたデータのみを用いて、時刻 t 4 でのパケット量の予測を行っても良い。更に、統計予測を行わずに、現在計測されたパケット量だけに基づいて負荷分散を行っても良い。

[0068]

ここで、統計情報の収集処理を、独立した機能部である統計情報収集処理部において行うことの利点を説明する。

[0069]

統計情報収集処理部15を独立した機能部とすることの利点は、装置コストの低減を図ることにある。通常のパケット転送を行う際には、パケット通信装置において統計情報収集処理は必ずしも必要がない。一方で、負荷分散等の高機能な処理を行う際においては、統計情報収集処理を行う必要が生じる。すなわち、パケット通信装置においては、統計情報収集処理は、付加機能であるといえる。したがって、統計情報収集処理部を独立に設けることにより、必要とするユーザに対してのみ統計情報収集処理を提供することが可能となり、装置コストの低減を図ることが可能である。

[0070]

さらに、収集する統計情報を変更する際など、統計情報収集処理部のみを変更 する(制御用ソフトウェアの変更、統計情報収集処理部を構成する機能モジュー ルの交換等)ことにより、対応が可能である。

[0071]

なお、パケット通信装置に複数の統計情報収集処理部を設け、回線対応部から

、同一のヘッダ転送フレームを、複数の統計情報収集処理部へ転送することにより、統計情報収集処理部を冗長化し、統計情報処理の負荷分散による処理能力向上や、耐障害性向上を図ることが可能である。

[0072]

ここまでの実施の形態では、パケット通信装置1における統計情報の収集処理 を、統計情報収集処理部15において行った上で、検索テーブル117を更新す る方法を説明した。

(実施例2)

統計情報の収集処理を、回線対応部11において行った上で、検索テーブル117を更新することも可能である。以下では、回線対応部11において統計情報の収集処理を行った上で、検索テーブル117を更新する方法の実施の形態を説明する。

[0073]

回線対応部11での統計情報の収集方法について説明する。

[0074]

図19は、回線対応部11において、統計情報を収集するための構成の一例を示す図である。パケット処理エンジン116は、受信パケットのヘッダ領域から、収集すべきフロー情報を読み出す。また、図示していないが、パケット処理エンジン116はパケットカウンタを備えており、ヘッダ情報を解析する機能とパケット量を計数する機能の両方を備えている。また、該受信パケットのバイト長、ビット長を計数する。データ加算部124は、パケット処理エンジン116が取り出したフローごとのパケット数、バイト数、ビット数を加算するためのレジスタを保持する。データ加算部124のレジスタに書き込まれたパケット数、バイト数、ビット数は、一定時間ごとに、図5で示した、統計情報収集処理部15に設ける統計テーブル154と同じフォーマットを持つ統計テーブル123に書き込む。

[0075]

検索テーブルの更新処理は、回線対応部11が行ってもよい。また、検索テーブルの更新処理は、回線対応部11で収集した統計情報を、制御部13に転送す

ることにより、制御部13で行ってもよい。

[0076]

回線対応部11において、検索テーブルの更新処理を行う場合は、図17、図18で示したフローチャートの手順にしたがって、プロセッサ121が処理を行う。

[0077]

検索テーブルの更新処理を、回線対応部が行う場合には、回線対応部11内で 処理を行うため、同一回線対応部が受信するパケットのフローのみを対象として 、負荷分散の振分先の更新を行う。

[0078]

検索テーブルの更新処理を制御部13が行う場合には、前記の統計情報収集処理部において検索テーブルの更新処理を行う場合と同様、パケット通信装置1が 受信する全てのパケットのフローを対象として、負荷分散の振分先の更新を行う

[0079]

検索テーブルの更新処理を制御部13が行う場合には、制御部13は、一定時間周期、たとえば15分間ごとに、装置内バス16を経由して、回線対応部11 の統計テーブル123を読み出すことにより、各回線対応部が収集した統計情報 を集約する。

[0080]

図20は、制御部13の構成の一例を示す図である。制御部13は、プロセッサ131、メモリ132、装置内バス16とのインタフェースである装置内バスインタフェース133、パケット通信装置1に対する制御装置とのインタフェースであるLANコントローラ134から構成される。

[0081]

制御部13のプロセッサ131は、各回線対応部11から転送された統計情報を、統合するテーブルを、メモリ132上に生成する。プロセッサ131は、前記のメモリ132上に生成したテーブルを元にして、図17のフローチャートによる手順と同様の処理を行うことにより、回線対応部11の検索テーブル117

の内容を更新する。制御部13は、前記の、統計情報収集処理部15から回線対応部11へのテーブル転送方法と同様の処理を行うことにより、回線対応部11 に対して、更新された検索テーブル117の内容を転送する。

[0082]

【発明の効果】

本発明のパケット通信装置によれば、負荷分散を行う上でのフローごとのパケットの振り分け先を、実際のトラヒックの状況に応じて、変更することが可能になる。

[0083]

また、本発明のパケット通信装置によれば、様々な利用形態に対応した統計情報収集処理を行うことが可能になる。

【図面の簡単な説明】

【図1】

本発明によるパケット通信装置の構成の一例を示す図。

【図2】

本発明によるパケット通信装置における回線対応部の構成の一例を示す図。

【図3】

本発明によるパケット通信装置における回線対応部に設ける検索テーブルの構成の一例を示す図。

【図4】

本発明によるパケット通信装置における統計情報収集処理部の構成の一例を示す図。

【図5】

本発明によるパケット通信装置における統計テーブルの構成の一例を示す図。

【図6】

本発明によるパケット通信装置を適用するネットワーク構成の一例を示す図。

【図7】

本発明によるパケット通信装置において、回線対応部からスイッチを経由して パケットを装置内で転送する際のフォーマットの一例を示す図。

【図8】

本発明によるパケット通信装置において、回線対応部から統計情報収集処理部 に対して、回線対応部が受信したパケットのヘッダ情報を転送する際のフレーム フォーマットの一例を示す図。

【図9】

本発明によるパケット通信装置において、回線対応部から統計情報収集処理部 に対して、回線対応部が受信したパケットのヘッダ情報を転送する際のフレーム フォーマットの一例を示す図。

【図10】

本発明によるパケット通信装置の回線対応部において、回線対応部が受信した パケットのヘッダ情報を統計情報収集処理部に対して転送する際の手順の一例を 示すフローチャート。

【図11】

本発明によるパケット通信装置の統計情報収集処理部において、統計情報を収集する手順の一例を示すフローチャート。

【図12】

Ethernetヘッダのフォーマットを示す図。

【図13】

IPヘッダ、TCPヘッダのフォーマットを示す図。

【図14】

VLANのフレームフォーマットを示す図。

【図15】

Ethernetを用いたMPLSのフレームフォーマットを示す図。

【図16】

RPRのヘッダフォーマットを示す図。

【図17]

本発明による統計情報収集処理部における検索テーブル更新手順の一例を示すフローチャート。

【図18】

本発明による統計情報収集処理部において、負荷分散の振分先を決定する手順の一例を示すフローチャート。

【図19】

本発明によるパケット通信装置における回線対応部の構成の一例を示す図。

【図20】

本発明によるパケット通信装置の制御部の構成の一例を示す図。

【図21】

本発明によるパケット通信装置における負荷分散の一例を示す図。

【図221

本発明によるパケット通信装置におけるヘッダ情報の転送方法の一例を示す図

【図23】

本発明によるパケット通信装置における検索テーブルを配布する方法の一例を示す図。

【図24】

本発明によるパケット通信装置における統計テーブルの構成の一例を示す図。

【図25】

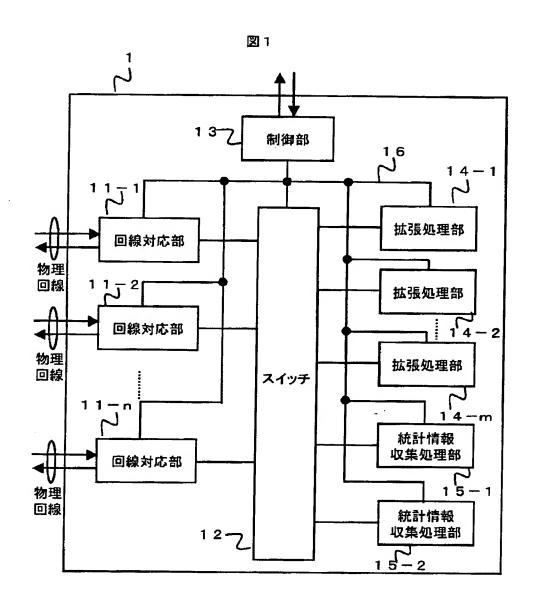
本発明によるパケット通信装置において、将来のパケット数の予測値を算出方法の一例を説明する図。

【符号の説明】

- 1 パケット通信装置
- 11 回線対応部
- 12 スイッチ
- 13 制御部
- 14 拡張処理部
- 15 統計情報収集処理部
- 16 装置内バス
- 30 装置内部パケット転送フォーマット
- 35 ヘッダ転送フレーム

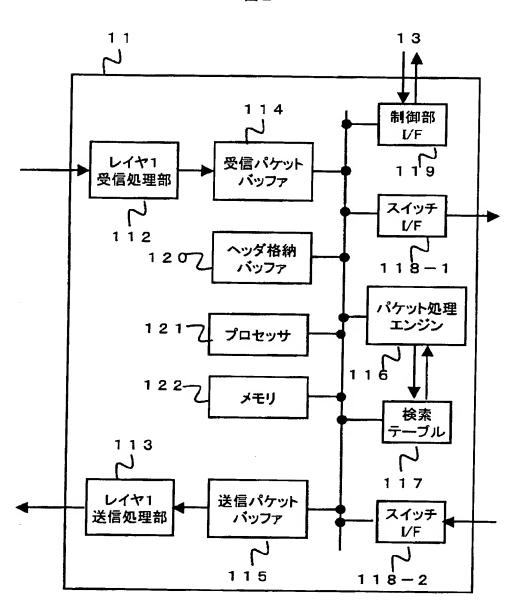
- 100 ネットワーク
- 117 検索テーブル
- 154 統計テーブル
- 310 装置内ヘッダ。

【書類名】 図面 【図1】



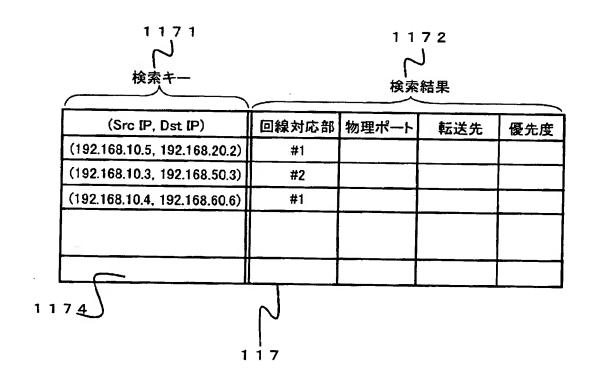
【図2】



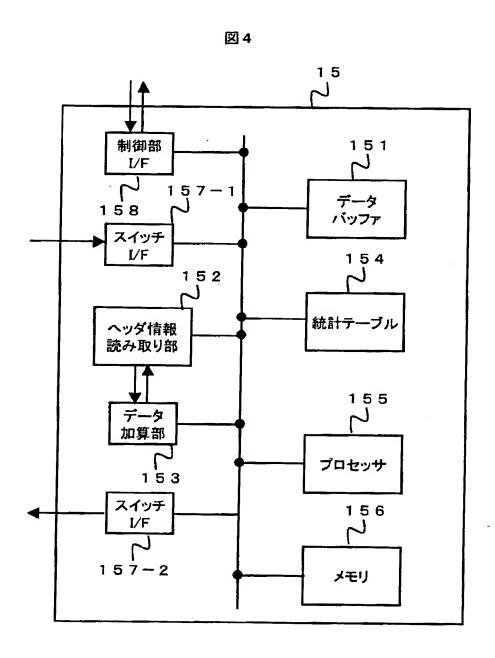


【図3】

図3

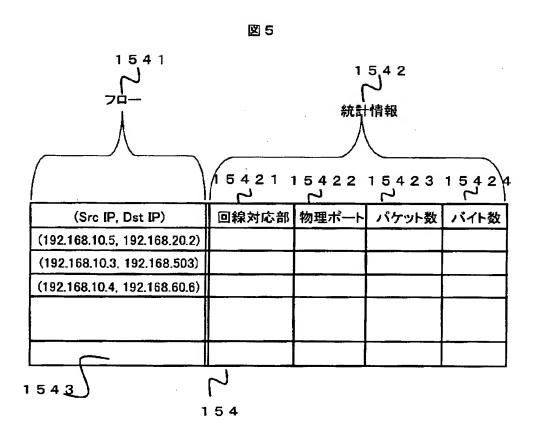


【図4】



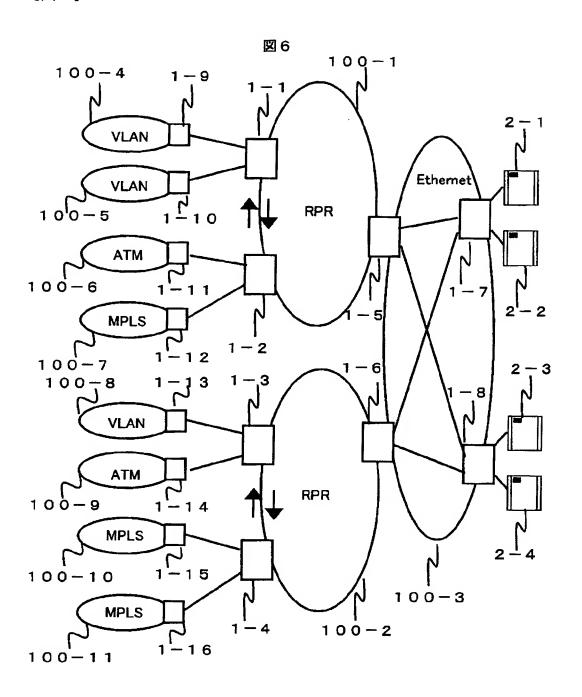


【図5】

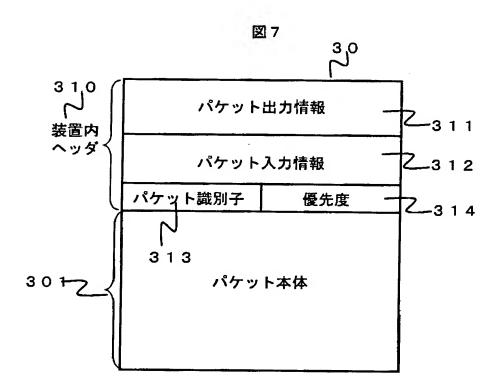




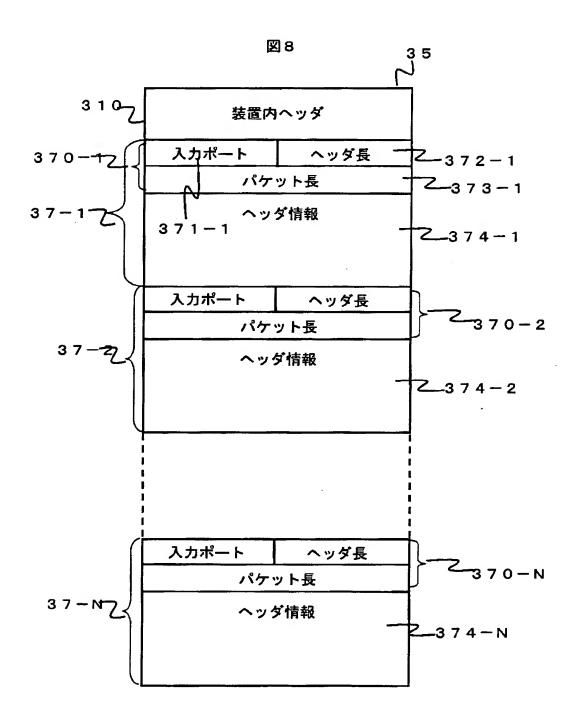
【図6】



【図7】

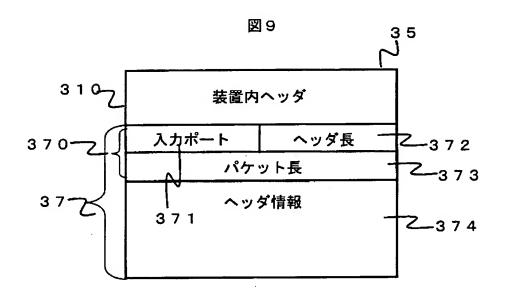


【図8】



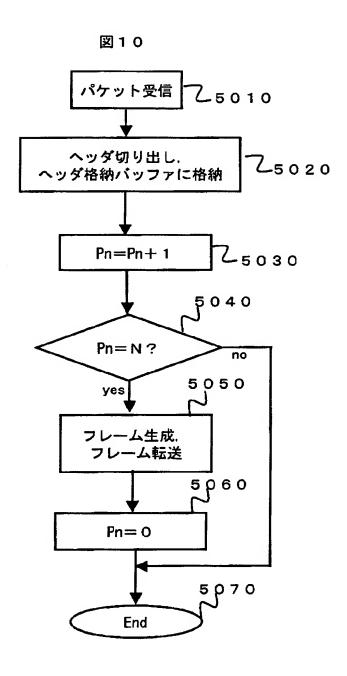


【図9】

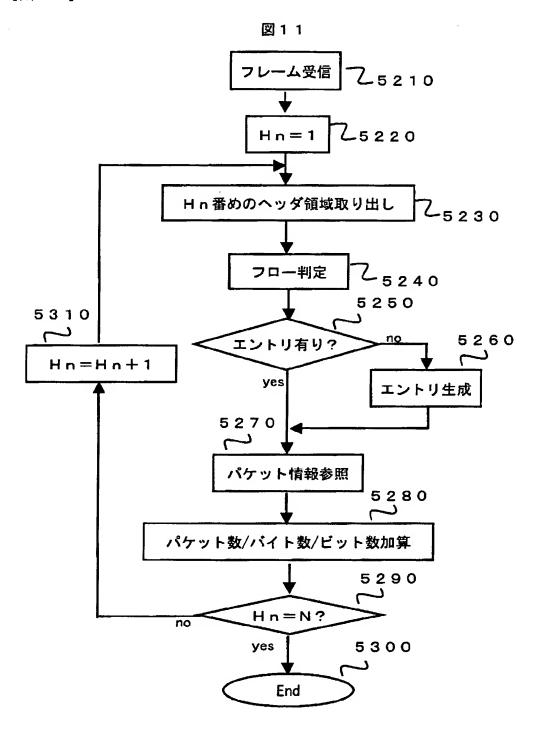


9

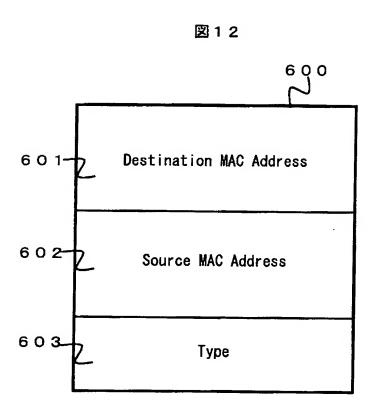
【図10】



【図11】

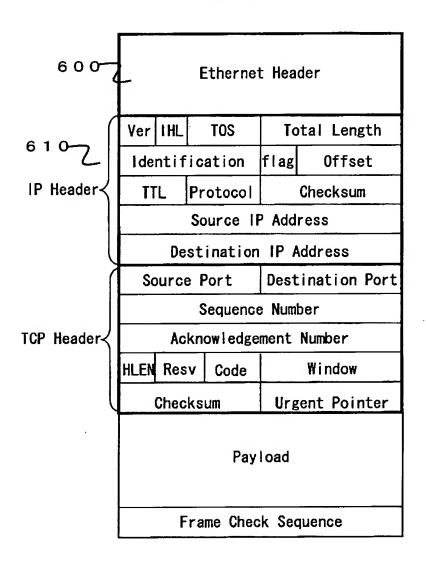


【図12】



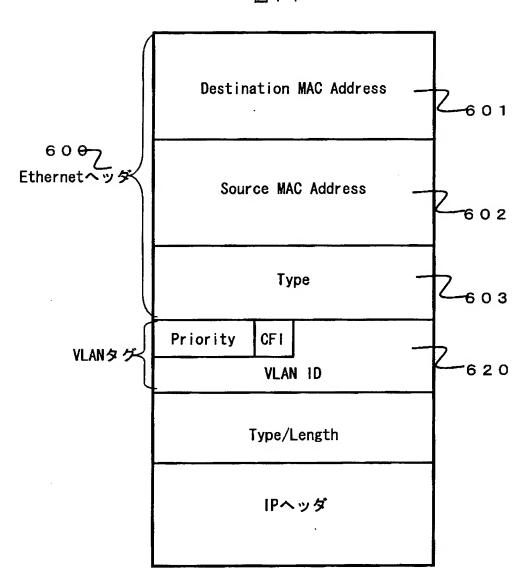
【図13】

図13

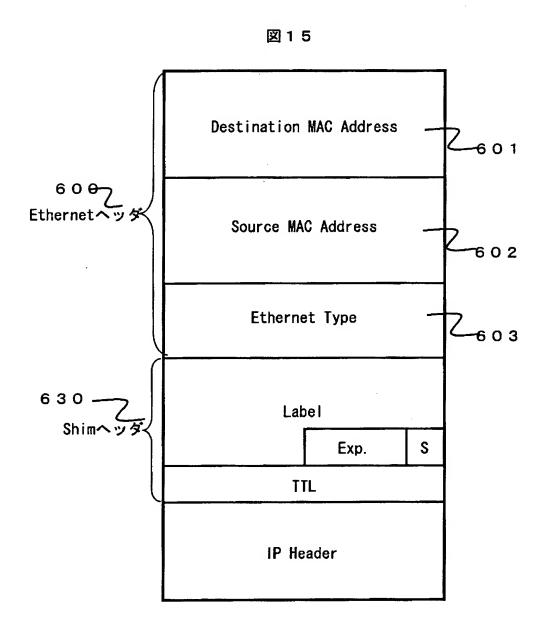


【図14】

図 1 4



【図15】



【図16】

TTL
RI FE PT SC WE R

Destination MAC Address

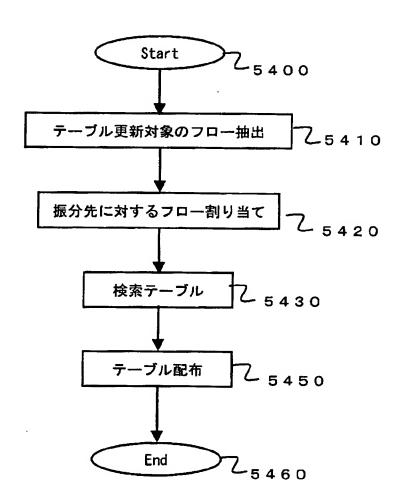
Source MAC Address

Ethernet Type

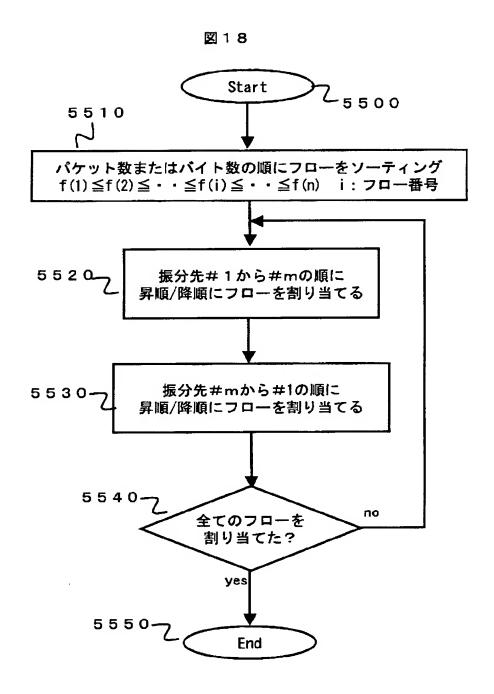
Header Checksum

【図17】

図17

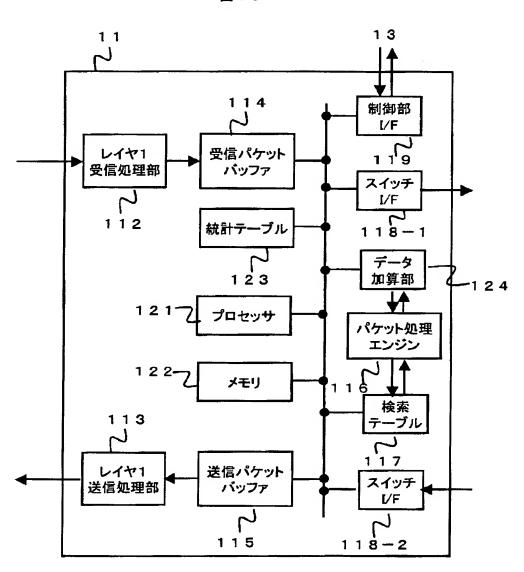


【図18】



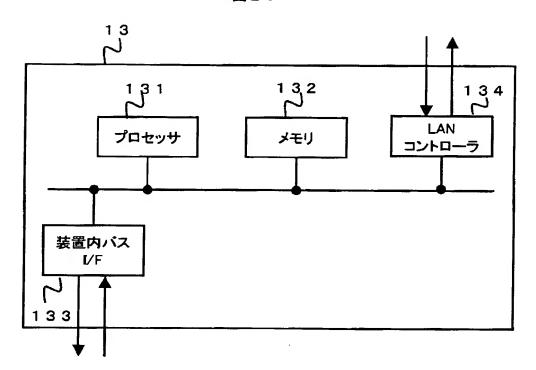
【図19】

図19

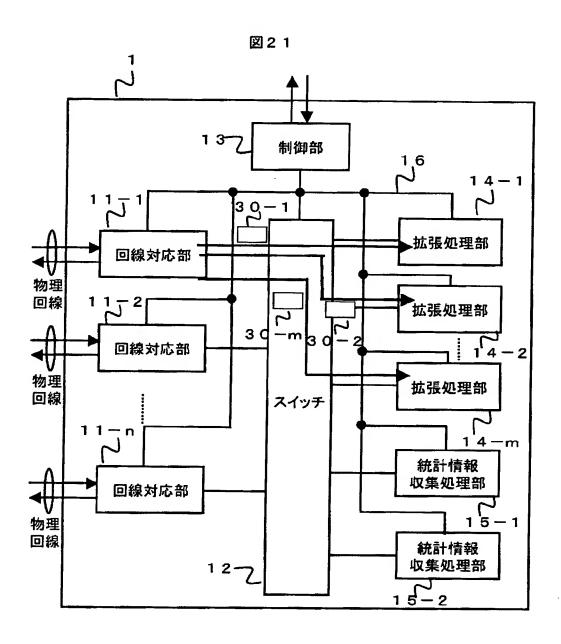


【図20】

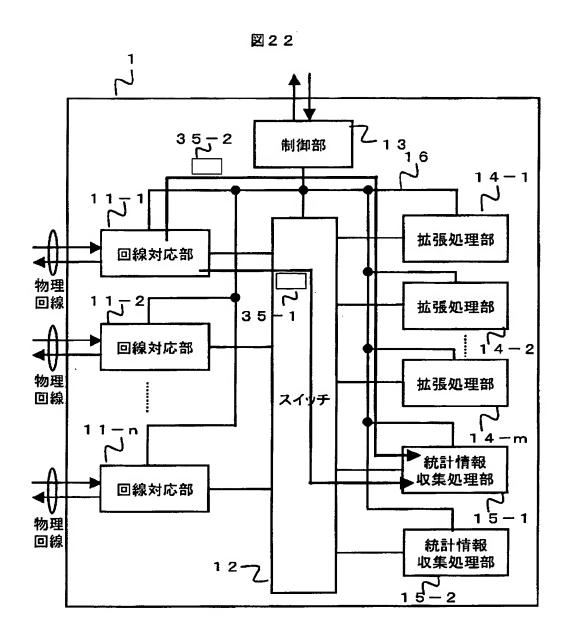




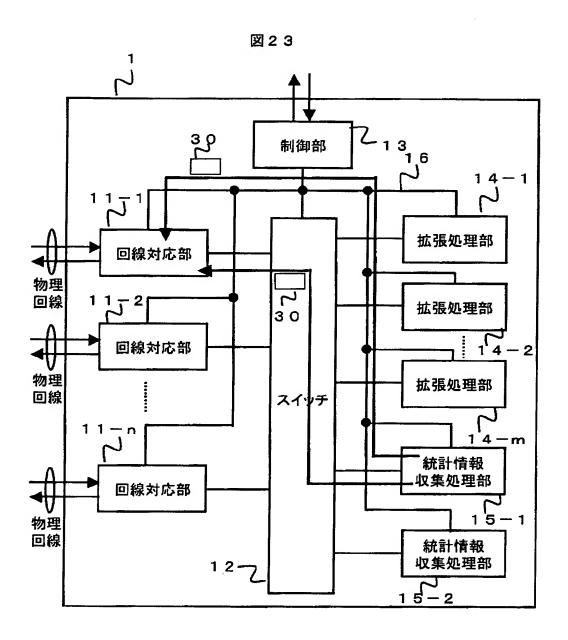
【図21】



【図22】

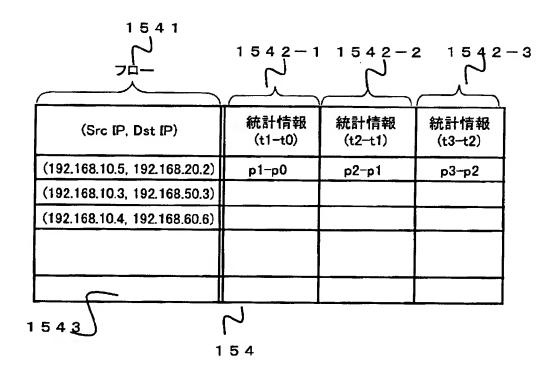


【図23】



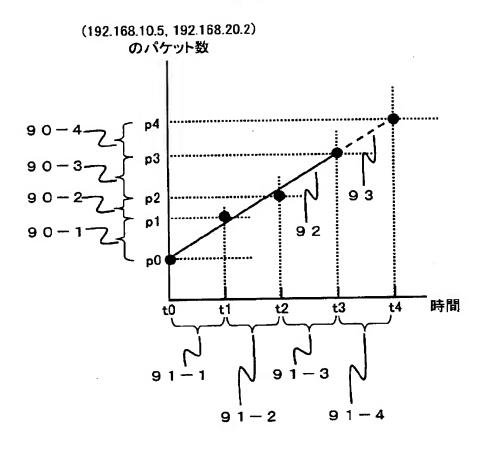
【図24】

図24



【図25】

図25



【書類名】

要約書

【要約】

【課題】 ネットワーク装置における負荷分散を行うのにあたり、トラヒック状況に即して、均等な振り分け処理を実現するパケット通信装置を提供する。

【解決手段】 パケット転送装置に、パケット転送処理とは独立した機能部として、統計情報収集処理部を設ける。回線対応部は、送受信するパケットのヘッダ情報を、統計情報収集処理部に転送する。統計情報収集処理部は、回線対応部から転送された送受信パケットのヘッダ情報により、統計情報を収集する。統計情報収集処理部が収集した統計情報に基づいて回線対応部に設ける検索テーブルの設定を更新する。

【選択図】 図1

認定・付加情報

特許出願の番号

特願2003-005237

受付番号

50300037827

書類名

特許願

担当官

第八担当上席 0097

作成日

平成15年 1月15日

<認定情報・付加情報>

【提出日】

平成15年 1月14日

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所